Current Zoology, 2016, 1–14
doi: 10.1093/cz/zow088
Advance Access Publication Date: 23 August 2016
Article

OXFORD

Article

Detecting polygenic selection in marine populations by combining population genomics and quantitative genetics approaches

Pierre-Alexandre Gagnaire[a,b,*] and Oscar E. Gaggiotti[c]

[a]Université Montpellier 2, Place Eugène Bataillon, 34095 Montpellier Cedex 5, France, [b]ISEM – CNRS, UMR 5554, SMEL, 2 rue des Chantiers, Sète, 34200, France, and [c]Scottish Oceans Institute, University of St Andrews, East Sands, St Andrews, KY16 9LB, UK

*Address correspondence to Pierre-Alexandre Gagnaire. E-mail: pagagnai@univ-montp2.fr.

Received on 14 March 2016; accepted on 21 July 2016

Abstract

Highly fecund marine species with dispersive life-history stages often display large population sizes and wide geographic distribution ranges. Consequently, they are expected to experience reduced genetic drift, efficient selection fueled by frequent adaptive mutations, and high migration loads. This has important consequences for understanding how local adaption proceeds in the sea. A key issue in this regard, relates to the genetic architecture underlying fitness traits. Theory predicts that adaptation may involve many genes but with a high variance in effect size. Therefore, the effect of selection on allele frequencies may be substantial for the largest effect size loci, but insignificant for small effect genes. In such a context, the performance of population genomic methods to unravel the genetic basis of adaptation depends on the fraction of adaptive genetic variance explained by the cumulative effect of outlier loci. Here, we address some methodological challenges associated with the detection of local adaptation using molecular approaches. We provide an overview of genome scan methods to detect selection, including those assuming complex demographic models that better describe spatial population structure. We then focus on quantitative genetics approaches that search for genotype–phenotype associations at different genomic scales, including genome-wide methods evaluating the cumulative effect of variants. We argue that the limited power of single locus tests can be alleviated by the use of polygenic scores to estimate the joint contribution of candidate variants to phenotypic variation.

Key words: local adaptation, genome scans, quantitative genetics, genotype-phenotype association, polygenic scores.

Introduction

Environmental changes, including those triggered by human activities, represent adaptive challenges to which natural populations may respond through changes in their genetic composition. In the sea as on land, understanding the genetic basis of the phenotypic changes underlying adaptation to environmental variation remains a fundamental objective. The application of next-generation sequencing technologies in ecological genomics has opened new perspectives for understanding how adaptation proceeds in nature (Davey et al. 2011; Savolainen et al. 2013). By providing genome-wide coverage of molecular variation in a large number of individuals, these technologies make it possible to use genetic methods that have long remained inaccessible to the study of non-model species. These methods broadly belong to 2 different approaches. One is based on a quantitative assessment of the total amount of genetic variation present for a given adaptive trait (Visscher et al. 2008). The second relies on the direct identification of genomic regions involved in adaptation, using genotype–phenotype correlations or selection tests (Stinchcombe and Hoekstra 2008). Both approaches have important practical implications in conservation biology for implementing molecular marker-based assessment of populations' adaptive potential

(Harrisson et al. 2014), especially for those facing climate change (Sgro et al. 2011; Franks and Hoffmann 2012), habitat modification (Schoville et al. 2012; Waits and Epps 2015), and overharvesting (Marty et al. 2015; Uusi-Heikkilä et al. 2015). However, despite these promising research avenues, the shift toward managing natural populations based on loci underlying adaptation to local conditions is still hindered by significant challenges (Eizaguirre and Baltazar-Soares 2014; McMahon et al. 2014). This is partly due to practical difficulties in quantifying the amount of genetic variation for adaptive traits in nature (Charmantier and Garant 2005), and identifying the genes that contribute to these traits (Rockman 2012).

The field of quantitative genetics specifically addresses the first part of this issue. If genetic variation is present for a given adaptive trait, a population's adaptive response (R) to a selection differential (S) can be predicted using the breeder's equation ($R = h^2S$). A central parameter in this equation is the narrow-sense heritability, which represents the proportion of the total phenotypic variance explained by the variance of additive genetic effects ($h^2 = V_A/V_P$) (Falconer and Mackay 1996; Lynch and Walsh 1998). This makes the estimation of heritability of prime importance for studying and predicting adaptive evolution (Visscher et al. 2008). However, determining the amount of genetic variation for fitness-related traits in nature remains very challenging. For instance, local adaptation can be confounded by plastic phenotypic responses to the environment, and heritability may differ among populations and environments (Charmantier and Garant 2005; Hansen et al. 2012). Fortunately, however, quantitative genetics approaches offer experimental designs that are specifically designed to address these issues, such as common garden or reciprocal transplants setups (Lynch and Walsh 1998; Kawecki and Ebert 2004).

The population genomic approach to local adaptation takes a different route. By focusing directly on the molecular signatures of selection, population genomics methods do not attempt to relate candidate loci to particular phenotypes. Thus, they can be implemented without a priori knowledge on the nature of adaptive traits (Luikart et al. 2003; Stinchcombe and Hoekstra 2008). Combined with the use of large molecular marker datasets, these so-called genome scan methods have accelerated the discovery of putatively adaptive variants in natural populations. However, the lists of candidate loci identified in genome scans are sometimes difficult to link with a particular mechanism of adaptation. Valuable insights into the selective agents can be obtained using methods that specifically search for associations between allele frequencies and environmental variables (e.g., Coop et al. 2010; de Villemereuil and Gaggiotti 2015). Ultimately, however, a more thorough understanding of the adaptive role played by candidate loci requires further assessments of their functional effects on phenotype (Storz and Wheat 2010; Barrett and Hoekstra 2011). Pioneering case studies have unraveled the different links connecting genotype to phenotype and fitness for simple traits governed by genes of large effect (e.g., Colosimo 2005; Linnen et al. 2009). However, fewer cases have been described for complex quantitative traits controlled by many genes of small effect (Hancock et al. 2011; Arnegard et al. 2014). More generally, our capacity to detect the genes that matter for adaptation in nature largely depends on the underlying genetic architecture of fitness-related traits, an important but usually unknown facet of adaptation.

Why does the genetic architecture of adaptive traits matter?

The genetic architecture refers to the number, genomic distribution, and effect sizes of genes that build and control a phenotypic trait and its variational constraints, including their mode of action such as additivity, dominance, epistasis, pleiotropy, and G × E interaction (Erickson et al. 2004; Hansen 2006).

A population's evolutionary response to environmental variation depends on the genetic architecture of adaptive traits. For instance, if adaptation involves several traits, the more genes that contribute to each trait, the more likely the response to selection will be affected by pleiotropic effects and linkage disequilibrium (LD) among genes. The resulting genetic correlations among traits may either increase or decrease the rate of adaptation, depending on the direction of maximum genetic variance relative to the optimum on the adaptive landscape (Lande 1979; Arnold 1992). Because these relationships among genes may be complex, a first important step is to understand whether adaptive traits are oligogenic or polygenic, that is, if they are essentially controlled by few genes of large effect or by large numbers of small effect genes.

Unfortunately, the evolution of genetic architecture is still poorly understood (Remington, 2015) due to a paucity of theoretical studies on this subject. Using a population genetic model for a trait under stabilizing selection, Rajon and Plotkin (2013) showed that the evolution of the genetic architecture depends on the strength of selection. Traits under either weak or strong selection are expected to be relatively oligogenic, whereas moderately selected traits should be encoded by many loci with a high variance in effect size. Thus, polygenic architectures are predicted mainly under intermediate selective intensities, which may correspond to the conditions encountered in large populations in which selection is efficient.

The effect of local adaptation in heterogeneous environments, with stabilizing selection favoring different optima in different populations connected by migration, tends to result in fewer loci of larger effects and tighter linkage (Yeaman and Whitlock 2011). This type of architecture is consistent with the view that large effect alleles better resist gene swamping when spatially varying selection is opposed by the homogenizing effect of gene flow (Lenormand 2002). However, local adaption can also occur by alleles of small effect that are prone to swamping, especially for traits that are genetically highly redundant (Yeaman 2015). In such cases, the genetic architecture is expected to be transient (i.e., there is a rapid turnover in locus contribution) and must be fueled by high mutation rates.

The genetic architecture of adaptive traits directly influences the response of individual quantitative trait loci (QTLs) to selection (Le Corre and Kremer 2012). A selected trait controlled by few genes of large effect should display strong allele frequency shifts at its underlying QTLs, whereas for a polygenic trait, a similar intensity of selection diluted among many small-effect genes should comparatively produce small allele frequency changes (Figure 1). Intermediate cases with many QTLs and a high variance in allele effect size (as predicted under moderate selection) should thus be characterized by a mixture of large and small allele frequency shifts, globally reflecting how the heterogeneous distribution of allele effect sizes is projected onto the genome. In such a situation, population genomic methods that detect signatures of selection by searching individual loci with outlying differentiation signals may be poorly adapted to study polygenic adaptation (Pritchard and Di Rienzo 2010). These general concerns, of course, also apply to the study of adaptation in marine species, but there are also theoretical and practical specificities that need to be considered for these species. Before reviewing the different methodologies offered by population genomics and quantitative genetics to detect selection, we present some key aspects that have important consequences for the study of local adaptation in marine species.

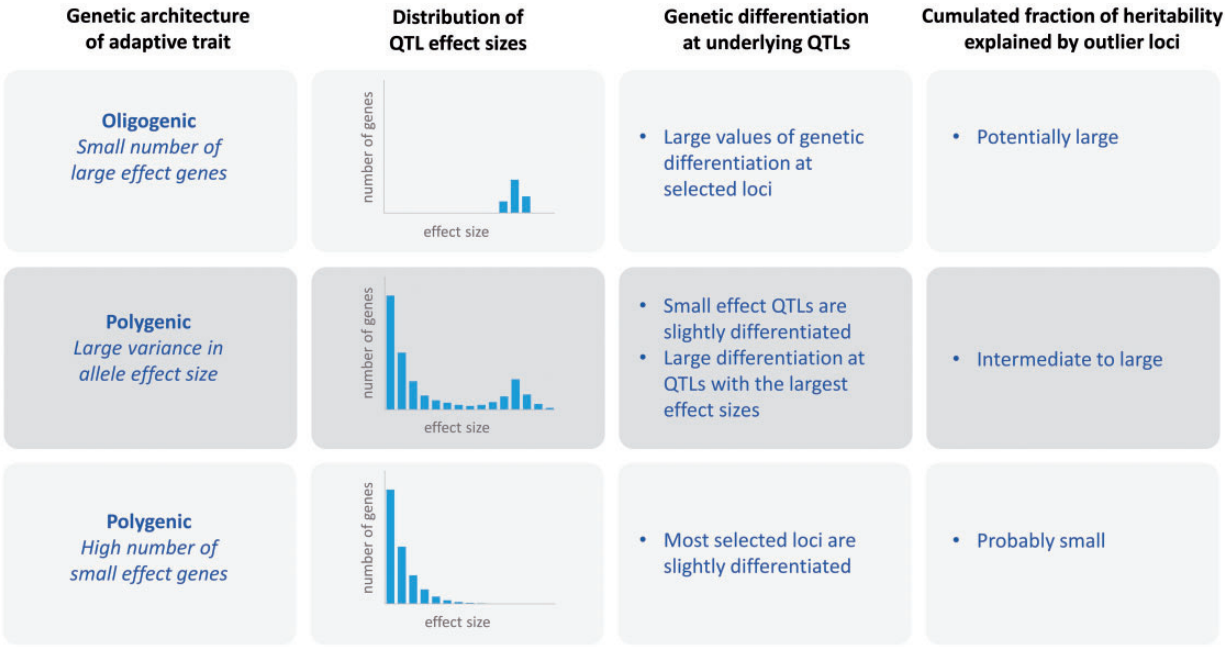

Figure 1. The consequence of the genetic architecture of an adaptive trait on our capacity to detect the molecular basis of local adaptation. Three different architectures are considered (oligogenic and polygenic with either large or small variance in allele effect size), together with a schematic representation of the distribution of allele effect sizes and their projection in terms of genetic differentiation at the underlying QTLs. Because population genomic methods detect selected loci based on their level of differentiation, the cumulated proportion of heritability explained by the joint contribution of outliers depends on the genetic architecture of adaptive traits.

Local adaptation in marine species: theoretical considerations

Marine species often challenge our perception of how geography and environment affect the genetic diversity of natural populations. In comparison to most terrestrial species, marine organisms usually display higher fecundity, larger population sizes, and higher dispersal potential, which generally result in weak to non-existent population genetic structure over broad spatial scales (Ward et al. 1994; Waples 1998; Palumbi 2003; Hedgecock et al. 2007). These peculiar life-history traits also have important consequences for the potential of marine populations to adapt to their environment. Because the efficiency of natural selection depends on population effective size, large populations are expected to have on average smaller proportions of effectively neutral mutations (i.e., those for which $s \ll 1/N_e$) (Ohta 1992). In theory, large populations are thus expected to be better adapted because even slightly advantageous mutations may contribute to adaptation, whereas at the same time, small effect deleterious mutations can be purged more efficiently (Eyre-Walker and Keightley 2007).

Increased efficiency of natural selection also has direct implications for adaption in a temporally or spatially varying environment. For instance, mutations that have opposite effects in different ecological contexts (antagonistic effects) or those that are selected in one environment while being neutral in others (conditional neutrality) can contribute to local adaptation only if they are visible to selection (i.e., if $s \gg 1/N_e$). In large populations, the evolutionary outcome of these genotype-by-environment interactions depends mostly on the relative strengths of migration and selection, with little influence of genetic drift (Slatkin 1973; Endler 1977). Thus, polymorphism can be durably maintained if a migration–selection equilibrium is reached, or lost by gene swamping if migration overwhelms the effect of selection (Bulmer 1972). Besides these variable outcomes of genotype-by-environment interactions, new locally beneficial alleles frequently appear in highly fecund species with large census sizes (Barton 2010). Therefore, widely dispersive and highly prolific marine species should retain high levels of adaptive genetic variation through both balancing selection and recurrent mutation (e.g., Schmidt and Rand 2001; Gagnaire et al. 2012; Pespeni and Palumbi 2013). A potential consequence of these theoretical predictions for high gene flow marine populations is that large-effect mutations that are swamping resistant should be found together with swamping-prone mutations of smaller effects that transiently contribute to local adaptation. This should result in polygenic architectures characterized by a high variance in allele effect size. Unfortunately, few empirical studies have examined the genetic architecture of adaptive traits in marine species.

Local adaptation in marine species: empirical evidence and limitations

The ideal experimental approach for detecting local adaptation, which consist in reciprocally transplanting genotypes among habitats (Kawecki and Ebert 2004), is practically impossible for most marine species. Therefore, tests for local adaptation more frequently involve the raising of individuals from different populations in a common experimental condition reproducing the main properties of a given habitat encountered in nature. Such common garden experiments have been employed in several marine species (reviewed in Sotka 2005; Conover et al. 2006; Sanford and Kelly 2011), including fishes (e.g., Hutchings et al. 2007; Hice et al. 2012), molluscs (e.g., Johannesson and Johannesson 1996), corals (e.g., Kenkel et al. 2015), and echinoderms (e.g., Pespeni et al. 2013). However, although they can be implemented more easily than reciprocal transplants, common garden experiments are still impracticable in many

broadcast spawning marine species, especially because it is often impossible to breed and rear progeny to maturity under laboratory conditions. These limitations also hinder the use of traditional QTL mapping approaches, which provide standard designs to statistically connect phenotypes to genotypes using experimental crosses.

Despite these practical difficulties, the study of local adaptation in marine organisms has a long history within the field of ecological population genetics (Williams et al. 1973; Powers and Place 1978; Koehn et al. 1980). Because of the aforementioned life-history characteristics, genetic variation patterns in many marine species are most often slightly influenced by spatially limited dispersal, although some physical oceanographic features may act as barriers to gene flow (Palumbi 1994; Hellberg 2009). Therefore, many marine species display genetic differentiation at locally adaptive loci, while being weakly (or almost not) differentiated at neutral markers (reviewed in Nielsen et al. 2009; Allendorf et al. 2010; Sanford and Kelly 2011; Gagnaire et al. 2015). Although weak neutral differentiation may only exist at small spatial scales (i.e., relative to species' ranges), the key point here is that environmental variation may occur at even finer scales. The resulting contrast in the differentiation level among markers is an undisputable advantage for applying genome scan methods that look for loci with exceptional levels of differentiation compared with the rest of the genome (de Villemereuil et al. 2014). On the downside, the chromosomal signature of local selection in high gene flow marine species is usually restricted to very small genomic regions (Gagnaire et al. 2015). Therefore, the detection of local adaptation loci often requires high-density genome scans. Recent studies that used this type of approach to detect loci influenced by selection have started to provide indication for both oligogenic and polygenic architectures (e.g., Pespeni et al. 2013; Bourret et al. 2014; Brieuc et al. 2015; Dixon et al. 2015; Hecht et al. 2015; Laporte et al. 2016). However, the variance in effect size was apparently large even when polygenic adaptation was supported.

If as suggested by theoretical and empirical works, quantitative fitness traits in marine species are encoded by many genes but with a high variance in allele effect size, then population genomic methods and genome-wide association studies (GWASs) should at least detect large and intermediate effect loci (Figure 1). On the other hand, the proportion of adaptive genetic variance that remains undetected by these approaches is generally unknown. This raises a 2-fold issue: (1) How can we improve the power to detect small effect loci and (2) how can we evaluate the joint contribution of candidate loci to variation in fitness traits? Below, we provide an overview of the current approaches that can be used to detect the genetic basis of local adaptation in marine populations and estimate the genetic contribution to variation in fitness-related traits. In the following sections, we consider the latest developments in population genomic methods to detect selection, focusing on the specific problem of detecting polygenic selection while avoiding false positives. We then focus on the quantitative genetics approaches based on molecular markers, especially those that can be implemented in natural populations with unrelated individuals.

Population Genomics Methods to Detect Polygenic Selection

The advent of next-generation sequencing (Shendure and Ji 2008) has made possible the generation of large datasets consisting of dense arrays of markers, typically single nucleotide polymorphisms (SNPs), covering the whole genome of a species. This in turn has led to a renewed interest in the development and application of statistical methods aimed at making inferences about the genetic architecture of adaptive traits. Given the inherent difficulties associated with the study of natural populations, and in particular marine species, the most popular methods adopt a population genomics approach based on collecting samples from natural populations and generating large number of molecular markers in order to scan the genome of species in search of so-called "outlier loci" whose behavior departs from the neutral expectations (Luikart et al. 2003; Storz 2005). Here we provide a short summary of the methods available and of the important difficulties associated with their use for the study of local adaptation in marine species.

Genome scan methods are meant to detect strong and moderate selection

New genome scan methods are constantly being developed but in general we can distinguish 2 main types of approaches: (1) based on allele frequencies, which implicitly or explicitly assume that loci are physically unlinked (for reviews see Narum and Hess 2011; Mita et al. 2013; de Villemereuil et al. 2014; Lotterhos and Whitlock 2014) and (2) based on the distribution of genetic variation along chromosomes, which take into account the physical linkage between markers (recently reviewed by Vatsiou et al. 2016). Here we will focus on the first type of methods, which is the best adapted to the study of most marine species, which are in general non-model organisms and lack extensive genetic resources such as a well-annotated genome and physical and genetic maps.

The simplest and most intuitive approach is based on analyzing the distribution of the allele frequency differential between 2 population samples (delta p). The rationale behind this approach is that the loci showing the largest allele frequency differential among populations are the most likely to be under selection. This seems particularly adapted to micro-geographic studies and selection experiments focused on single-generation selection footprints, where the initial allele frequencies are homogeneously distributed among sampling locations or experimental populations before selection starts. However, because the allele frequency change imposed by selection depends on the initial allele frequency before selection, large differences in allele frequency do not occur for low- or high-frequency variants even when selection is strong (Figure 2). Therefore, the tail of the distribution of delta p is enriched for common variants, whereas rare variants experiencing similar selection pressures cannot be detected. This justifies the use of more complex approaches that condition the level of genetic differentiation on allele frequencies or heterozygosity to detect outliers across the whole allele frequency spectrum.

The most popular approaches for the detection of selection are based on measures of genetic differentiation among populations, and can be traced back to Lewontin and Krakauer (1973). The underlying rationale is that divergent selection favoring different optimal phenotypes in different populations leads to strong genetic differentiation between them but only at the selected loci. Neutral loci, on the other hand are expected to exhibit much lower genetic differentiation because the homogenizing effect of migration is not counteracted by selection. Thus, it is possible to identify potentially selected genomic regions using locus-specific F_{ST} estimates, which are compared with either an empirical distribution (e.g., Akey 2002) or a distribution expected under neutrality (e.g., Beaumont and Nichols 1996; Beaumont and Balding 2004; Foll and Gaggiotti 2008; de Villemereuil and Gaggiotti 2015).

Another type of approach focused on allele frequency patterns looks for associations between environmental variables and allele frequencies at individual loci (e.g., Joost et al. 2007; Hancock et al.

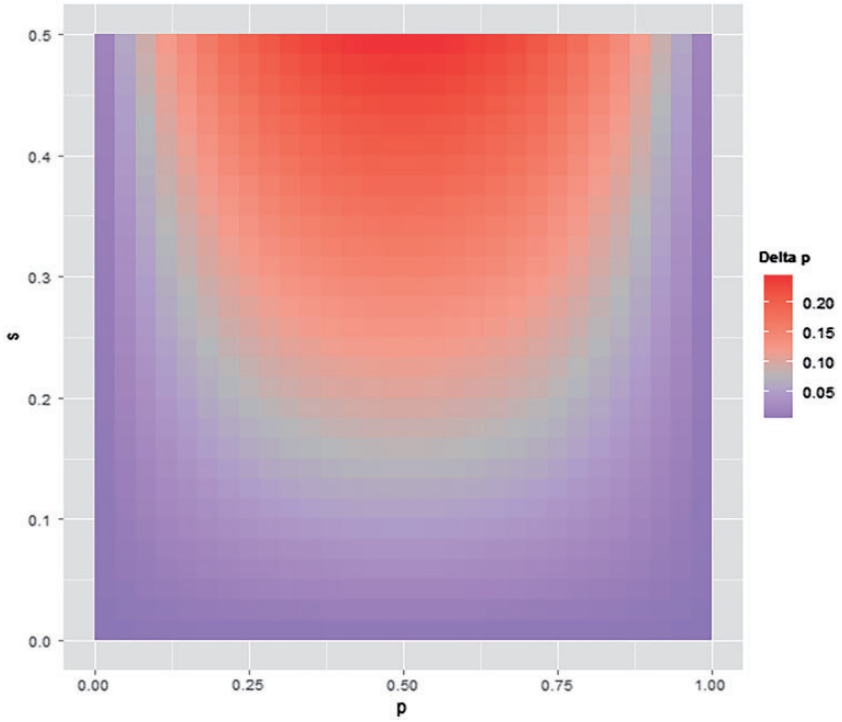

Figure 2. The effect of spatially varying selection in a symmetric additive viability model. A randomly mating population produces offspring that disperse randomly in 2 different habitats. After settlement in each habitat, the fitness of individual genotypes are $\omega_{AA}=1+s$, $\omega_{Aa}=1$, $\omega_{aa}=1-s$ in habitat 1, and $\omega_{AA}=1-s$, $\omega_{Aa}=1$, $\omega_{aa}=1+s$ in habitat 2, so that selection changes allele frequencies in opposite directions. The allele frequency differential measured between habitats after selection (delta *p*, colored scale) is shown as a function of the initial allele frequency before selection (*p*) and the strength of selection (*s*).

2008; Coop et al. 2010; Frichot et al. 2013; Guillot et al. 2014). The underlying rationale of these so-called "ecological association" methods (Frichot et al. 2015) is that in a heterogeneous habitat, environmental factors may exert a selective pressure for local adaptation. Thus, allele frequencies at loci underlying adaptive phenotypic traits should be associated with environmental factors that act as proxies for the unobserved selective pressures.

It should be noted that genome scan methods were not specifically developed to detect polygenic selection. Instead, they are grounded on the population genetic tradition of focusing on single-locus selection. Nevertheless, as mentioned before, theoretical and empirical works suggest that quantitative traits of marine species are encoded by many genes with a high variance in allele effect size. Thus, genome scans should be able to detect large and intermediate effect loci as suggested by a recent study that evaluated the performance of several methods under a scenario of polygenic selection (de Villemereuil et al. 2014). Moreover, in the case of species that have a well-annotated reference genome, it is possible to combine the output of genome scan methods with Gene Set Enrichment Analyses (Subramanian et al. 2005), to identify groups of genes that share a common biological function and underlie the metabolic pathway involved in local adaptation (e.g., Daub et al. 2013; Foll et al. 2014).

Accounting for more complex spatial structures and demographic effects

The application of genome scans to the study of local adaptation in natural populations involves important difficulties. Indeed, several evolutionary processes other than local adaptation can lead to genetic signatures similar to those left by selection. Probably the most discussed but still not completely resolved issue relates to demographic processes that create complex spatial patterns in allele and genotype frequencies. Riginos et al. (2016) discuss this problem in the general context of seascape genetics; here we focus on their effects on genome scan methods.

It is now well known that spatial population expansions can allow neutral alleles to reach very high frequency in newly colonized habitats by chance alone. The effects of this phenomenon, known as "allele-surfing" (Edmonds et al. 2004), can mimic a selective sweep (Excoffier et al. 2009a) and can lead to spurious associations between allele frequencies and environmental gradients and an inflation in the variance of F_{ST}. Hierarchical population structure, where local populations are grouped into regions or continents, can also increase the variance of F_{ST} at neutral loci over what is expected under a simple island model (Excoffier et al. 2009b). All this neutral demographic processes can lead to large false positive rates (FPRs). Two strategies have been proposed to try to account for these demographic effects. The first one adopted by genome scans based on allele frequency differentiation consists in explicitly assuming complex demographic models that better describe spatial population structure. For example, the approaches proposed by Excoffier et al. (2009b), Fariello et al. (2013), and Foll et al. (2014) explicitly assume a hierarchical island model. On the other hand, ecological association methods such as those proposed by Coop et al. (2010), Frichot et al. (2013), and Guillot et al. (2014) use phenomenological models that only seek to better describe the data using a statistical model that tries to capture the effects of unobserved demographic processes when estimating the effect of environmental factors on allele frequencies. Finally, a recent method, BayeScEnv (de Villemereuil and Gaggiotti 2015), extends the F_{ST}-based approach of Foll and Gaggiotti (2008) to include the locus-specific effect of an

environmental variable and a locus-specific effect that takes into account the biases due to violations to the island model assumed by the method.

The demographic processes mentioned above are applicable to any species. However, in the case of marine species we also need to consider the combined effects of ocean circulation processes and the high fecundity of species with larval dispersal; which could lead to large differences in reproductive success among individuals. This process, known as "sweepstakes reproductive success" (see Hedgecock and Pudovkin 2011 for a review), can generate chaotic genetic patchiness (Johnson and Black 1982; Broquet et al. 2013; Eldon et al. this issue); more importantly, it can cause F_{ST} to increase with migration rate (Yearsley et al. 2013). It is unclear whether or not these effects can lead to false signatures of selection such as allele surfing and inflation in the F_{ST} variance across loci. Recently Hoban et al. (2013) have shown that large variance in reproductive success can lead to negative Tajima's D (Tajima 1989) estimates, which could be erroneously interpreted as a signature of positive selection (or a bottleneck). However, spatial patterns generated by "sweepstakes reproductive success" are likely to be highly unstable and temporally variable, something that could limit the generation of false positives. Nevertheless, this issue needs to be investigated in more detail.

Finally, it is important to keep in mind that purely demographic processes are not the only source of false positives. More complex processes such as hybrid incompatibilities following secondary contact of diverged populations can generate strong LD (Kruuk et al. 1999) and spurious correlations between allele frequencies and environmental gradients (Barton and Hewitt 1985; Bierne et al. 2011). Also, purely genetic processes such as large differences in mutation rate across loci (Edelaar et al. 2011), and background selection (Charlesworth 1998) can increase FPR too. The only existing method that could in principle account for these additional biases is BayeScEnv (de Villemereuil and Gaggiotti 2015) but a more detailed sensitivity analyses of this method is needed to evaluate how it performs under these scenarios.

Combining Population Genomics and Quantitative Genetics Approaches

From F_{ST} genome scans to adaptive phenotypic variation

As mentioned earlier, genome scan methods search for genetic signatures of selection without explicitly considering the phenotypic traits involved in local adaptation. Therefore, population genomic approaches have the potential to simultaneously uncover the genetic basis of multiple traits that are jointly affected by selection. For example, this can happen along latitudinal gradients, where a suite of morphological, physiological, behavioral, and life-history traits of adaptive significance can be selected by various factors associated with latitude (Sanford and Kelly 2011; Hice et al. 2012). Without phenotypic information, genome scans cannot disentangle these multiple signals of selection and, therefore, are poorly adapted for uncovering the genetic architecture of local adaptation. A complementary approach that focuses on the phenotype itself is thus necessary to understand which traits are under selection (Merilä and Crnokrak 2001) and how variation in fitness-related traits can be linked with the candidate variants detected in F_{ST} genome scans (Stinchcombe and Hoekstra 2008; Barrett and Hoekstra 2011).

The quantitative genetics framework provides powerful designs which are especially well suited for studying quantitative traits in broadcast spawning marine species (Munday et al. 2013; Sunday et al. 2014; Davies et al. 2015). For instance, crossing experiments such as diallel and factorial breeding designs allow estimating the proportion of total phenotypic variance that is due to additive genetic variation, as well as the relative contributions of maternal and micro-environmental effects (Lynch and Walsh 1998). The approach advocated in a recent perspective on this subject relies on using common garden experiments to specifically deal with plasticity in genomic studies of local adaptation (de Villemereuil et al. 2016). This controlled experimental design can uncover the genetic component of phenotypic variation by standardizing the environment. Using appropriate replication for the different environmental conditions found in nature, common gardens also have the potential to reveal genotype-by-environment interactions. Quantitative genetics approaches have been widely used in aquaculture research to measure heritability, sometimes under different environmental conditions. These studies have provided evidence for the existence of genotype-by-environment interactions in several aquaculture species (Sae-Lim et al. 2015). Unfortunately, these designs have been rarely combined with molecular markers to identify QTLs that influence the variation of fitness-related traits in marine organisms.

Moderate-resolution linkage mapping can be conducted using experimental crosses (e.g., Colosimo et al. 2004; Gagnaire et al. 2013) or wild populations with known pedigrees (Slate et al. 2010). This approach relies on the co-segregation of known genetic markers with (unknown) neighboring QTL in genomic segments that are delimited by recent recombination events. It requires a relatively low density of markers and it involves low FPRs, although QTL studies with small sample sizes are sensitive to inflation of effect sizes due to the Beavis effect (Slate 2013). Linkage mapping needs information about individual pedigree, either under the form of a social pedigree established from field observations or family links identified using molecular markers. This type of information is unfortunately not available in large populations of unrelated individuals, which is a quite common situation in marine populations. A powerful alternative to detect QTL in populations that do not contain closely related individuals is to use a GWAS approach (Goddard and Hayes 2009). GWAS methods test for association between genetic markers and phenotypic variation on a SNP by SNP basis. Although they can be used to map QTLs with a much higher resolution than linkage mapping, they usually require a higher density of markers because recombination breaks down the statistical associations between QTL and neighboring markers over time. Large population genomic datasets now allow to conduct GWAS in non-model natural systems, providing invaluable information on the genetic architecture of important fitness traits, including the number, genomic distribution, effect size, and dominance patterns of QTL alleles (Barson et al. 2015). Other insightful studies have combined linkage mapping and GWAS approaches to evaluate their concordance (Santure et al. 2013, 2015), or have combined GWAS and F_{ST} genome scan approaches to identify genomic regions associated with variation in particular fitness traits (Johnston et al. 2014). This type of hybrid approach has been proposed to ascertain whether the candidate outlier loci detected in genome scans map to the same genomic regions as the QTL detected with phenotype-based methods (Stinchcombe and Hoekstra 2008; Barrett and Hoekstra 2011). Such a combination between population genomic and quantitative genetics methods is a promising avenue for understanding the complex links between phenotype, genotype, and fitness in the wild

(Munday et al. 2013; Jensen et al. 2014). However, in marine populations where LD is expected to be low, the success of these methods may require high marker densities and large sample sizes (at least several hundreds of individuals).

Another difficulty which is not limited to marine populations is that both genome scans and GWAS may be impaired by polygenic inheritance. High polygenicity requires the power to identify alleles whose effect is too weak to reach genome-wide or even nominal significance thresholds. Potentially, many small effect loci could collectively contribute to a non-negligible proportion of phenotypic variation. Thus, genome scans and GWAS may overlook an important fraction of the alleles that matter for evolution (Rockman 2012). In order to take this problem into account, some studies have proposed to use decreasing significance thresholds to estimate the cumulative effects of individual variants that are associated with phenotypic variation in GWAS (Purcell et al. 2009; Speliotes et al. 2010) (Figure 3A). For a given set of candidate loci, a polygenic score for cumulative effects can be obtained for each individual by either (1) summing the number of alleles that increase (or decrease) a trait value, or if available, (2) using allele effect sizes to weight the sum of alleles. To illustrate how polygenic scores are obtained, we take the example of a trait encoded by 5 loci and consider only alleles that increase the trait value. Furthermore, we use 0, 1, and 2 to denote the number of copies of the focal allele at each locus in a given individual. Then, a first individual with a combination of [1; 2; 0; 2; 1] allele counts at these loci would get an unweighted score of 6, whereas a second individual with genotype [0; 1; 1; 1; 2] would get a score of 5. Taking into account estimated allele effect sizes of [2, 1.5, 1.75, 1.25, 1.5] for these loci would give weighted scores of 9 and 7.5 for the first and second individual, respectively. Such individual polygenic scores can then be used to search for correlations with individual trait value or individual performance score (Arnegard et al. 2014). The same approach can be applied using candidate loci identified in genome scans. For instance, the sum of locally favorable alleles detected in a climate genome scan in *Arabidopsis thaliana* was shown to be a good predictor of individual local fitness, consistent with additive allelic effects across multiple loci (Hancock et al. 2011). These studies have paved the way for a more inclusive approach to local adaptation based on additive measures of individual polygenic scores. Lowering nominal significance thresholds to include additional small effect variants in the calculation of polygenic scores may be useful to estimate the minimal amount of phenotypic variance explained by the cumulative effect of candidate loci detected with genome scans. However, this approach also leads to an increase in FPRs. For the purpose of estimating the heritability explained by all the genetic markers together, alternative methods exist that do not search for individual QTL.

Alternatives to GWAS using pedigree-free approaches

The cumulative proportion of phenotypic variance explained by the significant QTLs detected by GWAS is usually less than the narrow-sense heritability estimated using pedigree information. This missing heritability has been attributed to the limited power to detect small effect QTL, but also to imperfect LD between the genotyped markers and the causative variants, especially if causal mutations segregate at lower frequencies than the marker loci (Yang et al. 2010). In order to avoid these effects, Yang et al. (2010) proposed a method for genome-wide complex trait analysis (GCTA) that fits all the markers simultaneously instead of searching for individual locus associations. This is done by using the genome-wide relatedness matrix (GRM) to estimate the additive genetic variance component within a linear mixed model (i.e., an animal model) fitted with a restricted maximum likelihood method. This approach is particularly well suited for large populations of unrelated individuals, and basically requires a large number of genetic markers genotyped in many samples to calculate the GRM containing all pairwise relatedness coefficients. Stanton-Geddes et al. (2013) determined that GCTA does not need to use the causative variants themselves if a sufficient number of markers is used to tag the causal variants. The method provided comparable h^2 estimates with either 25 000 SNPs (about 1 SNP per 10 kb) or 5 million SNPs, indicating that classical population genomics studies using GBS or RAD markers may reach fairly robust estimates of heritability. The minimal density of markers may, however, be higher in the presence of low LD. The sample size requirements may be of even greater concern, since basically the power of the method relies on the ability to use a large number of pairwise relatedness coefficients to fit the model. To illustrate this limitation, we used the distribution of pairwise relatedness coefficients obtained from 250 individuals sampled from a wild population of sea bream genotyped with c.a. 34 000 RAD SNPs (Figure 4). The sampled cohort mostly consists of unrelated specimens but also contains a few pairs of closely related individuals, and the resulting variance of pairwise SNP-derived genetic relationships is 0.00023. Using this empirical estimate of the variance in pairwise relatedness coefficients, we determined that more than 1300 individuals would be necessary to detect a non-null heritability ($h^2 > 0$) with a probability of at least 0.99 for a polygenic trait with a true heritability of 0.3. Moreover, the standard error of the estimate would be 0.07. Thus, genome-wide methods that estimate heritability in wild populations of unrelated individuals will usually require very large sample sizes to be reasonably powerful. Smaller sample sizes may, however, be sufficient for marine species with a sweepstake reproductive success (Hedgecock and Pudovkin 2011), because individuals from the same cohort may be strongly related to each other, but unrelated to other individuals sampled from a different cohort. Therefore, combining several cohorts should increase the variance of pairwise genetic relationships, thus providing increased power to detect heritability.

Another major interest of the quantitative genomics approach is that genome-wide relatedness coefficients can be estimated separately for each chromosome to partition the genetic variation for fitness traits across the genome (Yang et al. 2011; Robinson et al. 2013). Under the polygenic inheritance model, the proportion of phenotypic variance explained by each chromosome is proportional to its length or gene content (Figure 3B). This expectation provides a direct way to assess the hypothesis of polygenicity by evaluating the strength of the correlation between chromosome size and its contribution to overall additive genetic variance (Santure et al. 2013; Jensen et al. 2014). In practice, genome-wide and chromosome partitioning approaches will most often provide underestimates of the true heritability due to imperfect linkage between markers and causative variants. However, this effect should not affect the correlation expected between chromosome size and heritability for polygenic architectures.

High-density SNP datasets should enable even further partitioning of heritability by narrowing down the estimation of additive genetic variance to genomic segments within chromosomes (Figure 3C). This approach, called regional heritability mapping (Nagamine et al. 2012), has been recently implemented in a free-living population of Soay sheep to understand the genetic architecture of body size traits (Bérénos et al. 2015). Although no single SNP was found

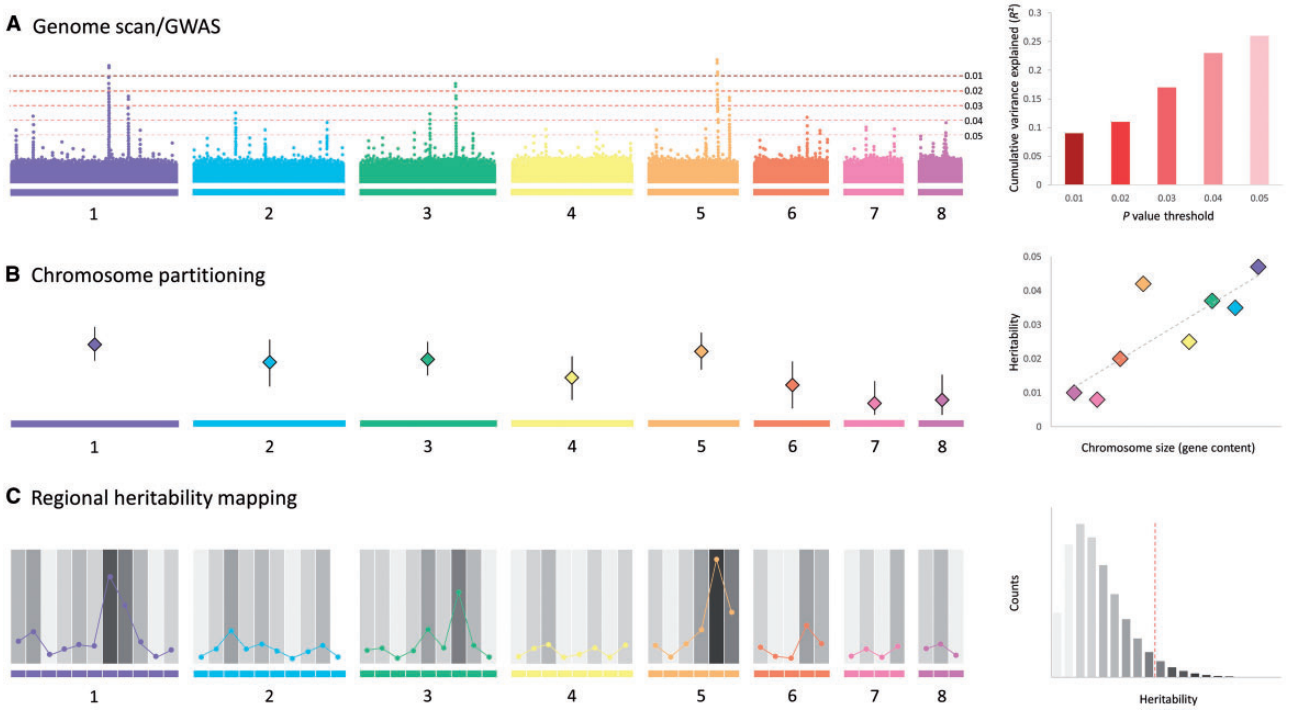

Figure 3. Conceptual plots illustrating 3 complementary approaches to dissect the genetic architecture of a complex fitness trait with a heritability of $h^2 = 0.3$. (**A**) In the genome scan/GWAS methodology, each locus (colored points on each of 8 chromosomes) is individually tested for genetic differentiation or association with the trait, which needs to control for multiple testing. The number of independent variants that are detected depends on the nominal detection threshold (horizontal dashed lines, $P = 0.05, 0.04, 0.03, 0.02, 0.01$). The proportion of phenotypic variance cumulatively explained by nominally significant variants increases with decreasing nominal significance thresholds (top right). (**B**) The quantitative genetics approach partitions the additive genetic variance across chromosomes by estimating relatedness using the SNPs present on each chromosome separately. Under the polygenic model of inheritance, the estimated heritability of each chromosome (colored diamonds with standard error bars) is positively correlated with its size or number of genes (middle right). In this ad hoc example, chromosome 5 explains more heritability than expected under the size/h^2 relationship, due to the presence of a large effect locus. (**C**) The same approach can be implemented by estimating heritability from multiple equal-sized genomic regions (connected dots in gray-shaded regions). Under the polygenic model, most regions contribute to a similar amount of heritability (light-gray-shaded areas). The largest effect loci likely reside within the regions that contribute disproportionately to phenotypic variance (dark-gray-shaded areas), lying above the 95th percentile of the distribution of regional heritability (bottom right, vertical dashed line) or explaining a significant amount of regional heritability in likelihood ratio tests.

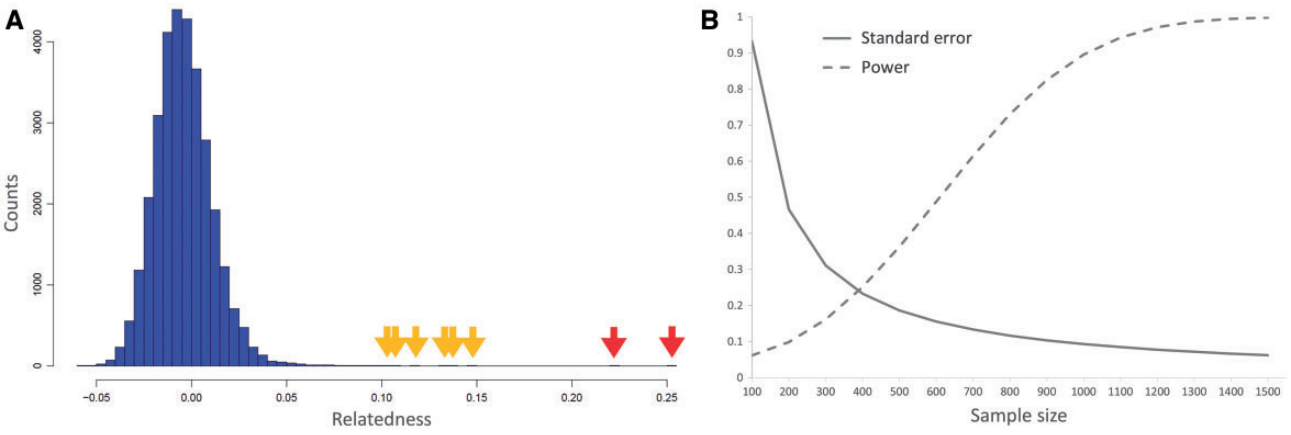

Figure 4. (A) Distribution of pairwise relatedness coefficients in a Mediterranean population of gilthead sea bream *Sparus aurata* from Southern France estimated using 34,000 SNPs (Gagnaire P-A, unpublished data). All samples are juveniles from the same cohort, red arrows indicate 2 half-sib pairs with relatedness coefficients close to 0.25, and orange arrows 6 pairs of individuals which are likely first cousins (relatedness coefficients close to 0.125). (**B**) The predicted standard error and power of SNP heritability estimate for a true heritability of $h^2 = 0.3$, given the observed variance of pairwise relatedness coefficients in the sea bream population sample (0.00023). Power is the probability of detecting $h^2 > 0$ given a type I error rate of 0.05. Standard error and power were estimated using GCTA-GREML Power Calculator (Visscher et al. 2014).

associated with these traits in GWAS, several regions of the Soay sheep genome contained SNPs that collectively explained significant amounts of phenotypic variance. This increased precision was obtained by fitting the GRM obtained with the SNPs present in each region together with the GRM calculated using all the remaining SNPs. In their study, Bérénos et al. (2015) used regions of 150 adjacent SNPs, although the same approach could be performed using equal-sized genomic fragments (Figure 3C). Genomic regions contributing significantly to heritability can be statistically detected using a likelihood ratio test that compare the likelihood of the genome-wide model with the likelihood of the genome-wide plus the regional model (Robinson et al. 2013; Bérénos et al. 2015). Because it represents a kind of intermediate between single marker and genome-wide heritability approaches, regional heritability mapping may help in identifying regions containing genes of too small effects to be individually detected. An additional strength of regional heritability mapping is that it should capture variation explained by rare variants in those regions.

The field of evolutionary genomics is now embracing methods and concepts that come from quantitative genetics, which combined with newly available genome-wide polymorphism data, will undoubtedly provide deeper insights into the genetic architecture of complex fitness traits. Therefore, the failure of the quantitative trait nucleotide program to discover adaptive mutations (Rockman 2012) may be partly overcome by combining population and quantitative genetics methods. For instance, it should be possible to estimate what proportion of the additive genetic variance in a given fitness-related trait can be explained by the outlier SNPs detected in genome scans for selection. This would not only provide a validation of the phenotypic effects of candidate mutations, but also a useful assessment of the missing heritability explained by the remaining (undetected) small effect mutations. Understanding and predicting local adaption in marine species could thus greatly benefit from this type of approach that attack the problem from both ends by focusing both on individual QTL and genome-wide effects on phenotypic variation. Importantly, the strategy sketched in Figure 3 can still be implemented even without a reference genome. In this case, chromosomal partitioning and regional heritability mapping would be of course impracticable, but the candidate loci identified in GWAS or genome scans can still be used to calculate polygenic scores, and pairwise relatedness coefficients can be calculated using genome-wide SNPs to estimate heritability.

Toward making genomic predictions

Genome-wide methods that estimate heritability from dense genotyping data use linear models to relate genetic to phenotypic variation. Similar approaches have been developed in animal breeding to estimate genomic breeding values (GEBV; Goddard and Hayes 2009). These genomic selection methods use a reference population which has been scored for phenotypes to derive an equation that predicts the breeding value of a given individual from its multilocus genotype. This prediction equation can then be applied to a new set of individuals which have been scored for genetic markers (but not necessarily for phenotypes) to predict their GEBV. High correlations between predicted GEBV and actual phenotypes have been obtained for highly heritable complex traits in a mouse population derived by crossbreeding inbred lines (Lee et al. 2008). These high prediction accuracies were probably facilitated by the high level of LD in the reference population. In marine species with large population sizes (where LD is supposed to be low), genomic prediction methods would probably need thousands of individuals and a high density of markers to be realistically applicable. However, the prediction accuracy which is required to address evolutionary questions is probably lower than the one needed in animal breeding or human medicine. Therefore, genomic prediction methods could in principle be implemented in wild populations to get estimates of unobserved phenotypes from genotype data (Jensen et al. 2014).

Applying genomic prediction methods in marine species would provide an invaluable way to address questions that cannot be answered due to practical difficulties to realize reciprocal transplants. For instance, what would be the phenotypic distribution of 2 populations adapted to 2 different environments if they were translocated to each other's environment? One way to answer this question would be to estimate allelic effects in each environment separately by generating a prediction equation relating individuals' genotypes to their phenotypes for each reference population in its native environment (Figure 5A). Then, the genetic value that each individual would have if it was translocated to the alternate environment can be derived using the prediction equation relating genotypes to phenotypes in the alternate environment. This kind of virtual reciprocal transplant would enable to predict what would be the phenotypic distribution of the genotypes sampled in environment 1 if they were living in environment 2, and vice versa. The same approach would also be useful to predict the potential phenotypic composition of a pool of pre-settled individuals if they were to settle in environment 1 or 2 (Figure 5B). Ultimately, it would help to understand how local selection shapes phenotypic diversity across environments by allowing the comparison of phenotypic diversity before (predicted) and after (observed) an episode of selection. Admittedly, such approach would be difficult to implement with more than 2 environments, but even so, it would still be useful to understand local adaption in some systems. Another potential limitation is that LD patterns should be roughly similar between the different populations to get reliable predictions, which mean that highly divergent populations are not ideal candidates to implement genomic prediction. Finally, trait heritability has to be sufficiently high for genetic values to be reasonably good predictors of the phenotype.

Another alternative to genomic prediction methods is to focus only on the candidate loci that are detected using genome scans or GWAS (Figure 3A). If a large enough proportion of phenotypic variance can be accounted for by variants that reach nominal significance, polygenic scores can be used to predict phenotypes. A methodological approach based on this idea has been developed by Berg and Coop (2014) to detect polygenic selection on quantitative traits from population genetic data. Their method estimates the mean additive genetic value for a phenotype in a given population using the significant SNPs detected in GWAS to compute the sum of allele frequencies weighted by the estimated allele effect sizes (i.e., a population equivalent of the individual polygenic score in Section 3.1). Then, they model the effect of genetic drift in a multipopulation model accounting for an arbitrary population structure to test for unusually strong correlations between genetic values and environmental variables. By looking for positive covariance of allelic effects, this approach has a greater power than single-locus approaches that focus on allele frequency changes. However, the applicability of this method for studies of marine populations remains to be evaluated especially with regards to the necessity to estimate allelic effects, because most population genomic studies still do not have enough power to estimate effect sizes correctly. Moreover, for genetic values to be treated as reliable phenotypic predictions across different environments, genotype-by-environment interactions should be taken into account as proposed for the genome-wide genomic prediction strategy (Figure 5).

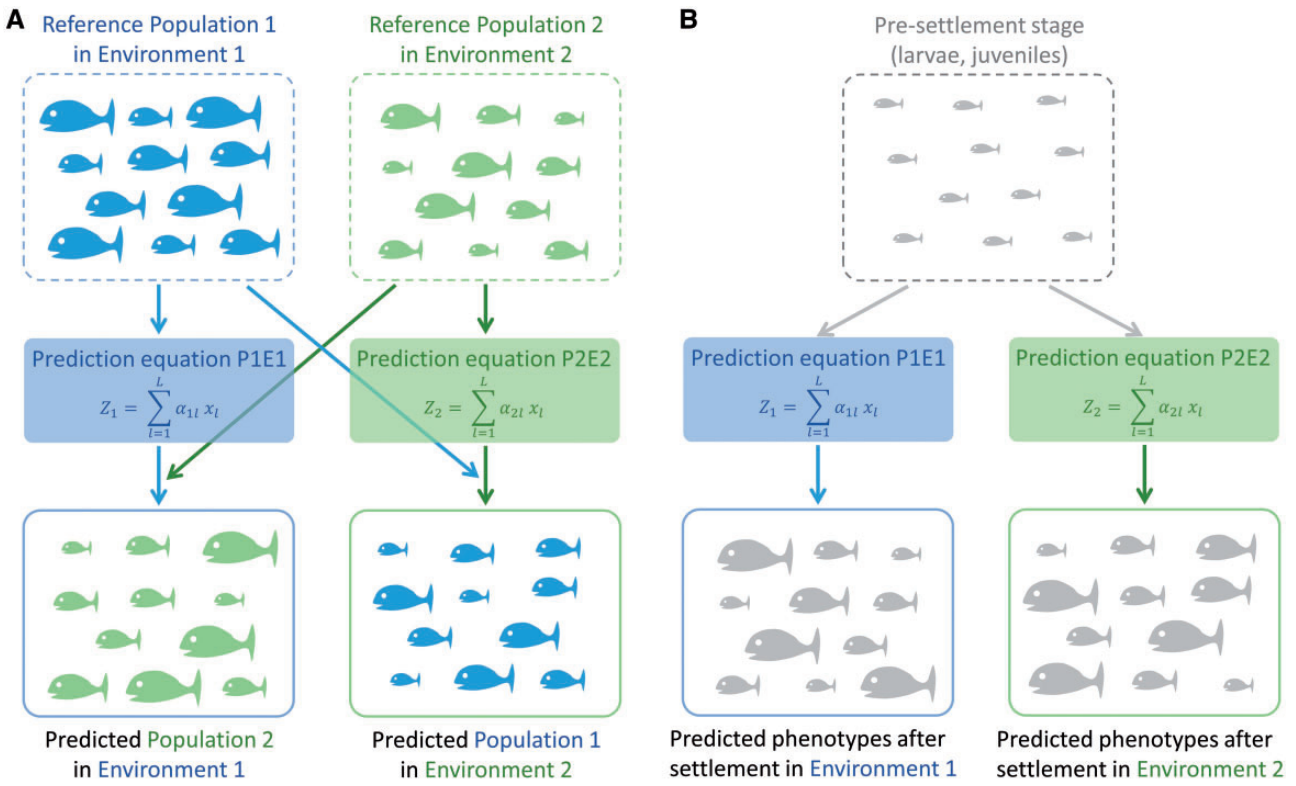

Figure 5. Using genomic prediction for making virtual transplants among different environments. (**A**) Two populations (P1 and P2) living in 2 different environments (E1 and E2) are genotyped and scored for a phenotypic trait (e.g., body size). In each population, a model predicting individual genetic value (Z_1 for P1E1, Z_2 for P2E2) is derived by combining the genotypes at all loci (for each locus, x is encoded additively as 0, 1, or 2 for *aa*, *aA* and *AA* genotypes, respectively) with their effects in the corresponding environment (α_{1l} in E1 or α_{2l} in E2). The prediction equation can then be applied to each reference population to predict the unobserved phenotype that each individual would have expressed in the alternate environment (P1 in E2 and P2 in E1). (**B**) The prediction equations can also be used to predict the eventual (unobserved) environment-dependent phenotypic composition of a group of young individuals sampled before settlement (e.g., larval or juvenile pool).

Conclusions

General guidelines

Practical considerations before diving into a genome-wide search for local adaptation in marine species include questions about marker density and sampling size. Ideally, the average distance of LD decay should be used to determine the minimal marker density required, such that virtually every SNP in the genome can be tagged by a genotyped marker. Less clear is the number of individuals needed to detect small signals of association and selection, especially under polygenic architectures. For instance, human studies which commonly use thousands of individuals are still underpowered to detect small-effect loci (Rockman 2012), raising important challenges for implementing comparable studies in non-model species. Genotyping costs can rapidly become prohibitive; therefore, depending on the available budget, a compromise has to be found between the density of markers and the number of individuals. However, with the progressive decrease in genotyping costs, we expect that studies in marine species using several hundreds to a few thousands samples will become common. Apart from these experimental design issues, 3 concluding guidelines can be considered:

i. First, local adaptation studies in marine species need a more systematic evaluation of the total amount of genetic variation in adaptive traits. Quantitative genetics provides a powerful framework to estimate heritability using specific experimental designs. However, when experimental crosses cannot be performed or when closely related individuals with known pedigree cannot be sampled, pedigree-free methods can be used to estimate heritability in the wild using measures of genome-wide relatedness among individuals.

ii. The second guideline concerns the issue of identifying genomic regions involved in fitness-related traits. We recommend to combine GWAS with genome scans for selection as much as possible in order to address genotype–phenotype and genotype-fitness links in parallel. Such an integrative approach has the potential to greatly improve our comprehension of the phenotypic effects of the many, but still often anonymous outliers detected in genome scan studies. Eventually, evaluating the relationship between individual locus effect size and genetic differentiation level will help to understand how the genetic architecture of phenotypic traits is projected onto the genome.

iii. We finally emphasize the need to estimate the joint contribution of the candidate loci detected by single locus methods. Polygenic scores can be used to adjust the detection thresholds in GWAS and genome scans in order to maximize the proportion of phenotypic variance explained by QTLs or outliers. Comparing this cumulative effect with the estimated value of heritability should make it possible to assess how much of genetic variance has been detected and, therefore, what is the remaining amount of missing heritability.

Looking forward

There is an increasing realization across the whole field of evolutionary biology, that a clear understanding of the local adaptation process and eventually speciation requires an integrative approach that explicitly considers the link between phenotype and genotype. For a very long time, the subfields of population and quantitative genetics progressed in parallel without much cross over but this situation is changing rapidly. In parallel with terrestrial studies, the field of marine ecological genomics is entering a new exciting phase that will benefit from a more systematic combination of population and quantitative genomic approaches. A successful integration of these 2 fields faces several empirical and methodological challenges that will need to be overcome. From an empirical point of view, this transition requires the ability to measure many phenotypic traits that could be involved in local adaptation in many individuals. In other words, we need to develop high-throughput phenotyping methods to complement NGS approaches. There are also statistical challenges requiring the development of new methods that can account for the complex spatial effects observed in marine species at both the genotype and phenotype level. Much work is being done in this regard in the field of statistical genetics but similar progress is needed in quantitative genetics and ultimately we need methods that integrate both types of data. These are important challenges but the potential rewards obtained from such an integrative approach are enormous.

A general solution that can overcome the above-mentioned challenges will take some time but in the meantime research projects that cannot implement this type of hybrid strategy will nevertheless benefit from the important recent methodological developments in both fields, which we outline in this review. Population genomic methods increasingly take into account the demographic complexity characterizing marine populations, and their power to detect polygenic selection is also increasing. On the other hand, pedigree-free approaches for estimating heritability open new possibilities for investigating the genetic basis of complex polygenic traits in natural populations. Despite these promising avenues of research, it should be kept in mind that any statistical method used to analyze samples from natural populations can only provide indirect evidence for the action of selection. This evidence always needs to be evaluated using experimental approaches that can confirm the findings of observational approaches. In other words, we should view the inferences drawn by observational approaches as hypotheses that need to be tested using sophisticated experimental approaches. An exciting prospect is the possibility of applying the methods we review here in the context of experimental settings as suggested by a recent review (de Villemereuil et al. 2016).

Acknowledgments

Cynthia Riginos and two anonymous reviewers provided very helpful comments that allowed us to improve the original manuscript.

Funding

P.-A.G. was supported by the CNRS-INEE action APEGE (*ArchiGen*). O.E.G. was supported by the Marine Alliance for Science and Technology for Scotland (MASTS).

References

Akey JM, 2002. Interrogating a high-density SNP map for signatures of natural selection. *Genome Res* **12**:1805–1814.

Allendorf FW, Hohenlohe PA, Luikart G, 2010. Genomics and the future of conservation genetics. *Nat Rev Genet* **11**:697–709.

Arnegard ME, McGee MD, Matthews B, Marchinko KB, Conte GL et al., 2014. Genetics of ecological divergence during speciation. *Nature* **511**:307–311.

Arnold SJ, 1992. Constraints on phenotypic evolution. *Am Nat* **140**:S85–S107.

Barrett RD, Hoekstra HE, 2011. Molecular spandrels: tests of adaptation at the genetic level. *Nat Rev Genet* **12**:767–780.

Barson NJ, Aykanat T, Hindar K, Baranski M, Bolstad GH et al., 2015. Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. *Nature* **528**:405–408.

Barton N, 2010. Understanding adaptation in large populations. *PLoS Genet* **6**:e1000987.

Barton NH, Hewitt GM, 1985. Analysis of hybrid zones. *Annu Rev Ecol Syst* **16**:113–148.

Beaumont MA, Balding DJ, 2004. Identifying adaptive genetic divergence among populations from genome scans. *Mol Ecol* **13**:969–980.

Beaumont MA, Nichols RA, 1996. Evaluating loci for use in the genetic analysis of population structure. *Proc R Soc B* **263**:1619–1626.

Bérénos C, Ellis PA, Pilkington JG, Lee SH, Gratten J et al., 2015. Heterogeneity of genetic architecture of body size traits in a free-living population. *Mol Ecol* **24**:1810–1830.

Berg JJ, Coop G, 2014. A population genetic signal of polygenic adaptation. *PLoS Genet* **10**:e1004412.

Bierne N, Welch J, Loire E, Bonhomme F, David P, 2011. The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Mol Ecol* **20**:2044–2072.

Bourret V, Dionne M, Bernatchez L, 2014. Detecting genotypic changes associated with selective mortality at sea in Atlantic salmon: polygenic multilocus analysis surpasses genome scan. *Mol Ecol* **23**:4444–4457.

Brieuc MSO, Ono K, Drinan DP, Naish KA, 2015. Integration of Random Forest with population-based outlier analyses provides insight on the genomic basis and evolution of run timing in Chinook salmon *Oncorhynchus tshawytscha*. *Mol Ecol* **24**:2729–2746.

Broquet T, Viard F, Yearsley JM, 2013. Genetic drift and collective dispersal can result in chaotic genetic patchiness. *Evolution* **67**:1660–1675.

Bulmer M, 1972. Multiple niche polymorphism. *Am Nat* **106**:254–257.

Charlesworth B, 1998. Measures of divergence between populations and the effect of forces that reduce variability. *Mol Biol Evol* **15**:538–543.

Charmantier A, Garant D, 2005. Environmental quality and evolutionary potential: lessons from wild populations. *Proc R Soc B* **272**:1415–1425.

Colosimo PF, 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* **307**:1928–1933.

Colosimo PF, Peichel CL, Nereng K, Blackman BK, Shapiro MD et al., 2004. The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biol* **2**:e109.

Conover DO, Clarke LM, Munch SB, Wagner GN, 2006. Spatial and temporal scales of adaptive divergence in marine fishes and the implications for conservation. *J Fish Biol* **69**:21–47.

Coop G, Witonsky D, Di Rienzo A, Pritchard JK, 2010. Using environmental correlations to identify loci underlying local adaptation. *Genetics* **185**:1411–1423.

Daub JT, Hofer T, Cutivet E, Dupanloup I, Quintana-Murci L et al., 2013. Evidence for polygenic adaptation to pathogens in the human genome. *Mol Biol Evol* **30**:1544–1558.

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM et al., 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat Rev Genet* **12**:499–510.

Davies SW, Scarpino SV, Pongwarin T, Scott J, Matz MV, 2015. Estimating Trait heritability in highly fecund species. *G3 Gene Genome Genet* **5**:2639–2645.

Dixon GB, Davies SW, Aglyamova GA, Meyer E, Bay LK et al., 2015. Genomic determinants of coral heat tolerance across latitudes. *Science* **348**:1460–1462.

Edelaar P, Burraco P, Gomez-Mestre I, 2011. Comparisons between QST and FST: how wrong have we been? *Mol Ecol* **20**:4830–4839.

Edmonds CA, Lillie AS, Cavalli-Sforza LL, 2004. Mutations arising in the wave front of an expanding population. *Proc Natl Acad Sci USA* **101**:975–979.

Eizaguirre C, Baltazar-Soares M, 2014. Evolutionary conservation: evaluating the adaptive potential of species. *Evol Appl* **7**:963–967.

Eldon B, Riquet F, Yearsley J, Jollivet D, Broquet T, 2016, this issue. Chaotic genetic patchiness. *Curr Zool* **62**.

Endler JA, 1977. *Geographic Variation, Speciation, and Clines*. Princeton (NJ): Princeton University Press.

Erickson DL, Fenster CB, Stenøien HK, Price D, 2004. Quantitative trait locus analyses and the study of evolutionary process. *Mol Ecol* **13**:2505–2522.

Excoffier L, Foll M, Petit RJ, 2009a. Genetic consequences of range expansions. *Annu Rev Ecol Evol Syst* **40**:481–501.

Excoffier L, Hofer T, Foll M, 2009b. Detecting loci under selection in a hierarchically structured population. *Heredity* **103**:285–298.

Eyre-Walker A, Keightley PD, 2007. The distribution of fitness effects of new mutations. *Nat Rev Genet* **8**:610–618.

Falconer DS, Mackay TF, 1996. *Introduction to Quantitative Genetics*. 4th edn. Harlow, Essex: Longmans Green.

Fariello MI, Boitard S, Naya H, SanCristobal M, Servin B, 2013. Detecting signatures of selection through haplotype differentiation among hierarchically structured populations. *Genetics* **193**:929–941.

Foll M, Gaggiotti O, 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a bayesian perspective. *Genetics* **180**:977–993.

Foll M, Gaggiotti Oscar E, Daub Josephine T, Vatsiou A, Excoffier L, 2014. Widespread signals of convergent adaptation to high altitude in Asia and America. *Am J Hum Genet* **95**:394–407.

Franks SJ, Hoffmann AA, 2012. Genetics of climate change adaptation. *Annu Rev Genet* **46**:185–208.

Frichot E, Schoville SD, Bouchard G, François O, 2013. Testing for associations between loci and environmental gradients using latent factor mixed models. *Mol Biol Evol* **30**:1687–1699.

Frichot E, Schoville SD, de Villemereuil P, Gaggiotti OE, François O, 2015. Detecting adaptive evolution based on association with ecological gradients: orientation matters & excl. *Heredity* **115**:22–28.

Gagnaire P-A, Broquet T, Aurelle D, Viard F, Souissi A et al., 2015. Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evol Appl* **8**:769–786.

Gagnaire P-A, Normandeau E, Côté C, Hansen MM, Bernatchez L, 2012. The genetic consequences of spatially varying selection in the panmictic American eel *Anguilla rostrata*. *Genetics* **190**:725–736.

Gagnaire P-A, Normandeau E, Pavey SA, Bernatchez L, 2013. Mapping phenotypic, expression and transmission ratio distortion QTL using RAD markers in the lake whitefish *Coregonus clupeaformis*. *Mol Ecol* **22**:3036–3048.

Goddard ME, Hayes BJ, 2009. Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nat Rev Genet* **10**:381–391.

Guillot G, Vitalis R, Le Rouzic A, Gautier M, 2014. Detecting correlation between allele frequencies and environmental variables as a signature of selection: a fast computational approach for genome-wide studies. *Spatial Stat* **8**:145–155.

Hancock AM, Brachi B, Faure N, Horton MW, Jarymowycz LB et al., 2011. Adaptation to climate across the *Arabidopsis thaliana* genome. *Science* **334**:83–86.

Hancock AM, Witonsky DB, Gordon AS, Eshel G, Pritchard JK et al., 2008. Adaptations to climate in candidate genes for common metabolic disorders. *PLoS Genet* **4**:e32.

Hansen MM, Olivieri I, Waller DM, Nielsen EE, 2012. Monitoring adaptive genetic responses to environmental change. *Mol Ecol* **21**:1311–1329.

Hansen TF, 2006. The evolution of genetic architecture. *Annu Rev Ecol Evol Syst* **37**:123–157.

Harrisson KA, Pavlova A, Telonis-Scott M, Sunnucks P, 2014. Using genomics to characterize evolutionary potential for conservation of wild populations. *Evol Appl* **7**:1008–1025.

Hecht BC, Matala AP, Hess JE, Narum SR, 2015. Environmental adaptation in Chinook salmon *Oncorhynchus tshawytsha* throughout their North American range. *Mol Ecol* **24**:5573–5595.

Hedgecock D, Barber PH, Edmands S, 2007. Genetic approaches to measuring connectivity. *Oceanography* **20**:70–79.

Hedgecock D, Pudovkin AI, 2011. Sweepstakes reproductive success in highly fecund marine fish and shellfish: a review and commentary. *Bull Mar Sci* **87**:971–1002.

Hellberg ME, 2009. Gene flow and isolation among populations of marine animals. *Annu Rev Ecol Evol Syst* **40**:291–310.

Hice LA, Duffy TA, Munch SB, Conover DO, 2012. Spatial scale and divergent patterns of variation in adapted traits in the ocean. *Ecol Lett* **15**:568–575.

Hoban SM, Mezzavilla M, Gaggiotti OE, Benazzo A, Van Oosterhout C et al., 2013. High variance in reproductive success generates a false signature of a genetic bottleneck in populations of constant size: a simulation study. *BMC Bioinform* **14**:309.

Hutchings JA, Swain DP, Rowe S, Eddington JD, Puvanendran V et al., 2007. Genetic variation in life-history reaction norms in a marine fish. *Proc R Soc B* **274**:1693–1699.

Jensen H, Szulkin M, Slate J, 2014. Molecular quantitative genetics. In: Charmantier A, Garant D, Kruuk LEB, editors. *Quantitative Genetics in the Wild*. Oxford: Oxford University Press, 209–227.

Johannesson B, Johannesson K, 1996. Population differences in behaviour and morphology in the snail *Littorina saxatilis*: phenotypic plasticity or genetic differentiation? *J Zool* **240**:475–493.

Johnson M, Black R, 1982. Chaotic genetic patchiness in an intertidal limpet *Siphonaria* sp. *Mar Biol* **70**:157–164.

Johnston SE, Orell P, Pritchard VL, Kent MP, Lien S et al., 2014. Genome-wide SNP analysis reveals a genetic basis for sea-age variation in a wild population of Atlantic salmon *Salmo salar*. *Mol Ecol* **23**:3452–3468.

Joost S, Bonin A, Bruford MW, DesprÉS L, Conord C et al., 2007. A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Mol Ecol* **16**:3955–3969.

Kawecki TJ, Ebert D, 2004. Conceptual issues in local adaptation. *Ecol Lett* **7**:1225–1241.

Kenkel C, Setta S, Matz M, 2015. Heritable differences in fitness-related traits among populations of the mustard hill coral *Porites astreoides*. *Heredity*. **115**:509–516.

Koehn RK, Newell RI, Immermann F, 1980. Maintenance of an aminopeptidase allele frequency cline by natural selection. *Proc Natl Acad Sci* **77**:5385–5389.

Kruuk LEB, Baird SJE, Gale KS, Barton NH, 1999. A comparison of multilocus clines maintained by environmental adaptation or by selection against hybrids. *Genetics* **153**:1959–1971.

Lande R, 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution* **33**:402–416.

Laporte M, Pavey S, Rougeux C, Pierron F, Lauzent M et al., 2016. RAD sequencing reveals within-generation polygenic selection in response to anthropogenic organic and metal contamination in North Atlantic Eels. *Mol Ecol* **25**:219–237.

Le Corre V, Kremer A, 2012. The genetic differentiation at quantitative trait loci under local adaptation. *Mol Ecol* **21**:1548–1566.

Lee SH, van der Werf JH, Hayes BJ, Goddard ME, Visscher PM, 2008. Predicting unobserved phenotypes for complex traits from whole-genome SNP data. *PLoS Genet* **4**:e1000231.

Lenormand T, 2002. Gene flow and the limits to natural selection. *Trends Ecol Evol* **17**:183–189.

Lewontin RC, Krakauer J, 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics* **74**:175–195.

Linnen CR, Kingsley EP, Jensen JD, Hoekstra HE, 2009. On the origin and spread of an adaptive allele in deer mice. *Science* **325**:1095–1098.

Lotterhos KE, Whitlock MC, 2014. Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Mol Ecol* **23**:2178–2192.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P, 2003. The power and promise of population genomics: from genotyping to genome typing. *Nat Rev Genet* **4**:981–994.

Lynch M, Walsh B, 1998. *Genetics and Analysis of Quantitative Traits*. Sunderland (MA): Sinauer.

Marty L, Dieckmann U, Ernande B, 2015. Fisheries-induced neutral and adaptive evolution in exploited fish populations and consequences for their adaptive potential. *Evol Appl* **8**:47–63.

McMahon BJ, Teeling EC, Höglund J, 2014. How and why should we implement genomics into conservation? *Evol Appl* **7**:999–1007.

Merilä J, Crnokrak P, 2001. Comparison of genetic differentiation at marker loci and quantitative traits. *J Evol Biol* **14**:892–903.

Mita S, Thuillet AC, Gay L, Ahmadi N, Manel S et al., 2013. Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Mol Ecol* **22**:1383–1399.

Munday PL, Warner RR, Monro K, Pandolfi JM, Marshall DJ, 2013. Predicting evolutionary responses to climate change in the sea. *Ecol Lett* **16**:1488–1500.

Nagamine Y, Pong-Wong R, Navarro P, Vitart V, Hayward C et al., 2012. Localising loci underlying complex trait variation using regional genomic relationship mapping. *PLoS ONE* **7**:e46501.

Narum SR, Hess JE, 2011. Comparison of FST outlier tests for SNP loci under selection. *Mol Ecol Resour* **11**:184–194.

Nielsen EE, Hemmer-Hansen J, Larsen PF, Bekkevold D, 2009. Population genomics of marine fishes: identifying adaptive variation in space and time. *Mol Ecol* **18**:3128–3150.

Ohta T, 1992. The nearly neutral theory of molecular evolution. *Annu Rev Ecol Syst* **23**:263–286.

Palumbi SR, 1994. Genetic divergence, reproductive isolation, and marine speciation. *Annu Rev Ecol Syst* **25**:547–572.

Palumbi SR, 2003. Population genetics, demographic connectivity, and the design of marine reserves. *Ecol Appl* **13**:S146–S158.

Pespeni MH, Palumbi SR, 2013. Signals of selection in outlier loci in a widely dispersing species across an environmental mosaic. *Mol Ecol* **22**:3580–3597.

Pespeni MH, Sanford E, Gaylord B, Hill TM, Hosfelt JD et al., 2013. Evolutionary change during experimental ocean acidification. *Proc Natl Acad Sci* **110**:6937–6942.

Powers DA, Place AR, 1978. Biochemical genetics of *Fundulus heteroclitus* (L.). I. Temporal and spatial variation in gene frequencies of Ldh-B, Mdh-A, Gpi-B, and Pgm-A. *Biochem Genet* **16**:593–607.

Pritchard JK, Di Rienzo A, 2010. Adaptation: not by sweeps alone. *Nat Rev Genet* **11**:665–667.

Purcell SM, Wray NR, Stone JL, Visscher PM, O'donovan MC et al., 2009. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**:748–752.

Rajon E, Plotkin JB, 2013. The evolution of genetic architectures underlying quantitative traits. *Proc R Soc B* **280**:20131552.

Remington DL, 2015. Alleles versus mutations: understanding the evolution of genetic architecture requires a molecular perspective on allelic origins. *Evolution* **69**:3025–3038.

Riginos C, Crandall ED, Liggins L, Bongaerts P, Treml EA, 2016, this issue. Navigating the currents of seascape genomics: how spatial analyses can augment population genomic studies. *Curr Zool* **62**.

Robinson MR, Santure AW, DeCauwer I, Sheldon BC, Slate J, 2013. Partitioning of genetic variation across the genome using multimarker methods in a wild bird population. *Mol Ecol* **22**:3963–3980.

Rockman MV, 2012. The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution* **66**:1–17.

Sae-Lim P, Gjerde B, Nielsen HM, Mulder H, Kause A, 2015. A review of genotype-by-environment interaction and micro-environmental sensitivity in aquaculture species. *Rev Aquacult*.

Sanford E, Kelly MW, 2011. Local adaptation in marine invertebrates. *Annu Rev Mar Sci* **3**:509–535.

Santure AW, De Cauwer I, Robinson MR, Poissant J, Sheldon BC et al., 2013. Genomic dissection of variation in clutch size and egg mass in a wild great tit *Parus major* population. *Mol Ecol* **22**:3949–3962.

Santure AW, Poissant J, De Cauwer I, van Oers K, Robinson MR et al., 2015. Replicated analysis of the genetic architecture of quantitative traits in two wild great tit populations. *Mol Ecol* **24**:6148–6162.

Savolainen O, Lascoux M, Merilä J, 2013. Ecological genomics of local adaptation. *Nat Rev Genet* **14**:807–820.

Schmidt PS, Rand DM, 2001. Adaptive maintenance of genetic polymorphism in an intertidal barnacle: habitat-and-life-stage specific survivorship of MPI genotypes. *Evolution* **55**:1336–1344.

Schoville SD, Bonin A, François O, Lobreaux S, Melodelima C et al., 2012. Adaptive genetic variation on the landscape: methods and cases. *Annu Rev Ecol Evol Syst* **43**:23–43.

Sgro CM, Lowe AJ, Hoffmann AA, 2011. Building evolutionary resilience for conserving biodiversity under climate change. *Evol Appl* **4**:326–337.

Shendure J, Ji H, 2008. Next-generation DNA sequencing. *Nat Biotech* **26**:1135–1145.

Slate J, 2013. From Beavis to beak color: a simulation study to examine how much QTL mapping can reveal about the genetic architecture of quantitative traits. *Evolution* **67**:1251–1262.

Slate J, Santure AW, Feulner PGD, Brown EA, Ball AD et al., 2010. Genome mapping in intensively studied wild vertebrate populations. *Trends Genet* **26**:275–284.

Slatkin M, 1973. Gene flow and selection in a cline. *Genetics* **75**:733–756.

Sotka EE, 2005. Local adaptation in host use among marine invertebrates. *Ecol Lett* **8**:448–459.

Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G et al., 2010. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet* **42**:937–948.

Stanton-Geddes J, Yoder JB, Briskine R, Young ND, Tiffin P, 2013. Estimating heritability using genomic data. *Methods Ecol Evol* **4**:1151–1158.

Stinchcombe JR, Hoekstra HE, 2008. Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity* **100**:158–170.

Storz JF, 2005. Using genome scans of DNA polymorphism to infer adaptive population divergence. *Mol Ecol* **14**:671–688.

Storz JF, Wheat CW, 2010. Integrating evolutionary and functional approaches to infer adaptation at specific loci. *Evolution* **64**:2489–2509.

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL et al., 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci* **102**:15545–15550.

Sunday JM, Calosi P, Dupont S, Munday PL, Stillman JH et al., 2014. Evolution in an acidifying ocean. *Trends Ecol Evol* **29**:117–125.

Tajima F, 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**:585–595.

Uusi-Heikkilä S, Whiteley AR, Kuparinen A, Matsumura S, Venturelli PA et al., 2015. The evolutionary legacy of size-selective harvesting extends from genes to populations. *Evol Appl* **8**:597–620.

Vatsiou AI, Bazin E, Gaggiotti OE, 2016. Detection of selective sweeps in structured populations: a comparison of recent methods. *Mol Ecol* **25**:89–103.

de Villemereuil P, Frichot É, Bazin É, François O, Gaggiotti OE, 2014. Genome scan methods against more complex models: when and how much should we trust them? *Mol Ecol* **23**:2006–2019.

de Villemereuil P, Gaggiotti O, Mouterde M, Till-Bottraud I, 2016. Common garden experiments in the genomic era: new perspectives and opportunities. *Heredity* **116**:249–254.

de Villemereuil P, Gaggiotti OE, 2015. A new FST-based method to uncover local adaptation using environmental variables. *Methods Ecol Evol* **6**:1248–1258.

Visscher PM, Hemani G, Vinkhuyzen AA, Chen G-B, Lee SH et al., 2014. Statistical power to detect genetic (co) variance of complex traits using SNP data in unrelated samples. *PLoS Genet* **10**:e1004269.

Visscher PM, Hill WG, Wray NR, 2008. Heritability in the genomics era: concepts and misconceptions. *Nat Rev Genet* **9**:255–266.

Waits LP, Epps CW, 2015. Population genetics and wildlife habitat. *Wildlife Habitat Conserv* **63**.

Waples RS, 1998. Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *J Hered* **89**:438–450.

Ward RD, Woodwark M, Skibinski DOF, 1994. A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes. *J Fish Biol* **44**:213–232.

Williams GC, Koehn RK, Mitton JB, 1973. Genetic differentiation without isolation in the American eel *Anguilla rostrata*. *Evolution* **27**:192–204.

Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK et al., 2010. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* **42**:565–569.

Yang J, Manolio TA, Pasquale LR, Boerwinkle E, Caporaso N et al., 2011. Genome partitioning of genetic variation for complex traits using common SNPs. *Nat Genet* **43**:519–525.

Yeaman S, 2015. Local adaptation by alleles of small effect. *Am Nat* **186**:S74–S89.

Yeaman S, Whitlock MC, 2011. The genetic architecture of adaptation under migration-selection balance. *Evolution* **65**:1897–1911.

Yearsley JM, Viard F, Broquet T, 2013. The effect of collective dispersal on the genetic structure of a subdivided population. *Evolution* **67**:1649–1659.